The Sampling Distribution of the Sample Mean

Suppose that a variable x of a population has mean, μ and standard deviation, Then, for samples of size n,

1) The mean of $\bar{x}$ equals the population mean, μ, in other words: $\mu\bar{x} = \mu$

2) The standard deviation of $\bar{x}$ equals the population standard deviation divided by the square root of the sample size, in other words: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

3) If x is normally distributed, so is $\bar{x}$, regardless of sample size

4) If the sample size is large ($n > 30$), $\bar{x}$ is approximately normally distributed, regardless of the distribution of x.

Examples:

1) The times that college students spend studying per week have a distribution that is right skewed with a mean of 8.4 hours and a standard deviation of 2.7 hours. Suppose a random sample of 45 students is selected.

a) What is the sampling distribution of the mean number of hours these 45 students spend studying per week? (i.e., What is the sampling distribution of $\bar{x}$?)

$n =$ _____ $\mu =$ _____ $\sigma =$ _____

b) Find the probability that the mean time spent studying per week is between 8 and 9 hours. Find the probability that the mean time spent studying per week is greater than 9.5 hours.

2) At an urban hospital the weights of newborn babies are normally distributed, with a mean of 7.2 pounds and standard deviation of 1.2 pounds. Suppose a random sample of 30 is selected.

a) What is the sampling distribution of the mean weight of these newborn babies? (i.e. What is the sampling distribution of $\bar{x}$?)

b) Find the probability the mean weight is less than 6.9 pounds?

c) Find the probability the mean weight is between 6.5 and 7.5 pounds?

d) Find the probability the mean weight is greater than 8 pounds?

3) A battery manufacturer claims that the lifetime of a certain battery has a mean of 40 hours and a standard deviation of 5 hours. A simple random sample of 100 batteries is selected.

a) What is the sampling distribution of the mean life of the batteries?

(i.e. What is the sampling distribution of $\bar{x}$?)

b) What is the probability the mean life is less than 38.5? Would this be unusual?

WEIGHTED MEANS AND MEANS AS WEIGHTED SUMS

In the Speeds Problem we saw that there is more than one kind of "average." In this handout, we will explore this topic further. The ordinary mean is sometimes called the "arithmetic mean" to distinguish it from other types of means.

** The most common way to think of the average (arithmetic mean) of numbers is to add them up and divide by the total number of summands:

e.g., the average of $1,1,2,3,4,4,4$ *is* $(1+1+2+3+4+4+4)/7$

But we could write these two other ways:

1. **"Distributing" the denominator gives**

$$(1/7)1 + (1/7)1 + (1/7)2 + (1/7)3 + (1/7)4 + (1/7)4 + (1/7)4.$$

Thus, we have the mean as a sum of coefficients times the original numbers in the list. *Note that the sum of the coefficients is 1.*

2. **Collecting like terms gives**

$$(2x1 + 2 + 3 + 3x4)/7 = (2/7)1 + (1/7)2 + (1/7)3 + (3/7)4.$$

Now we have a sum of coefficients times the *distinct* values (not allowing repetitions) in the original list of numbers. The coefficient of a value is the *proportion* of that value in the original list of numbers. We still have the coefficients adding to 1, but they are no longer all the same. We now see the mean as a *weighted sum* of the *distinct values*, where each value is weighted according to its proportion in the total list of numbers. This perspective prompts two generalizations of the arithmetic mean.

A. **Weighted Means**

To form a *weighted mean* of numbers, we first multiply each number by a number ("weight") for that number, then add up all the weighted numbers, then divide by the sum of the weights. We often do this in computing course